Necesarry libraries

```
library(readr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```
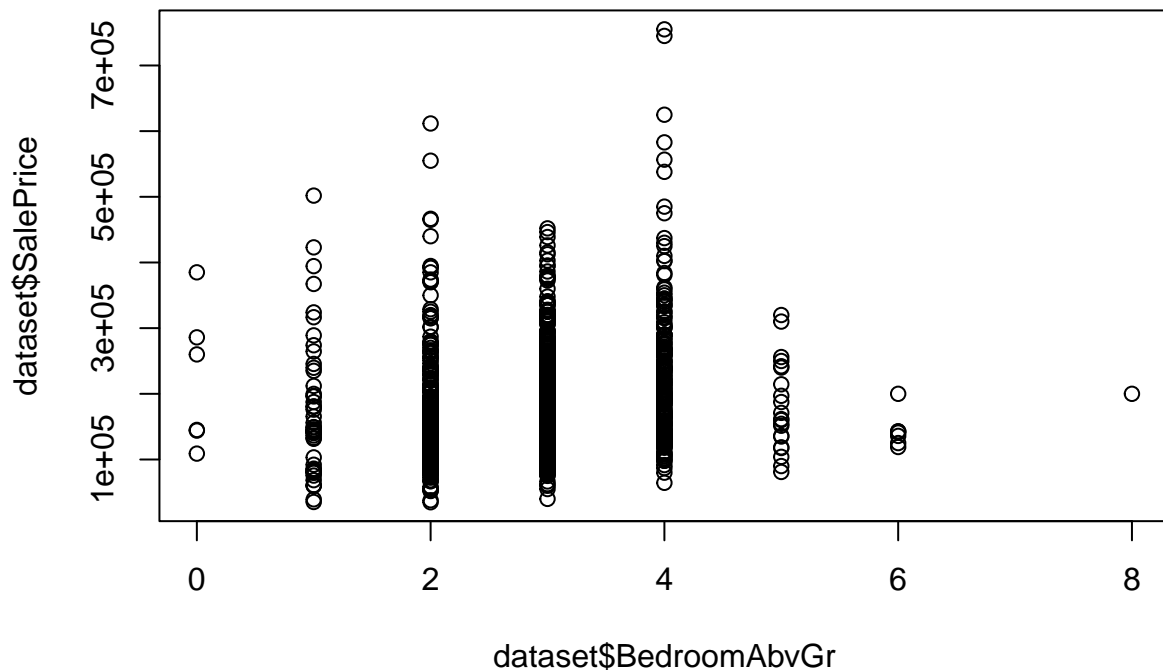
Loading dataset

```
dataset <- read_csv("preprocessed_data.csv")
```

```
## Rows: 1460 Columns: 81
## -- Column specification ------------------------------------------------------
## Delimiter: ","
## chr (43): MSZoning, Street, Alley, LotShape, LandContour, Utilities, LotConf...
## dbl (38): Id, MSSubClass, LotFrontage, LotArea, OverallQual, OverallCond, Ye...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
regression = lm(BedroomAbvGr ~ SalePrice,data=dataset)

plot(dataset$BedroomAbvGr,dataset$SalePrice)
lines(dataset$BedroomAbvGr,regression$fitted.values,col='red')
```

dataset

```
## # A tibble: 1,460 x 81
##        Id MSSubClass MSZon~1 LotFr~2 LotArea Street Alley LotSh~3 LandC~4 Utili~5
##     <dbl>      <dbl> <chr>     <dbl>   <dbl> <chr>  <chr> <chr>   <chr>   <chr>
## 1      1         60 RL           65    8450 Pave   <NA>  Reg     Lvl     AllPub
## 2      2         20 RL           80    9600 Pave   <NA>  Reg     Lvl     AllPub
## 3      3         60 RL           68   11250 Pave   <NA>  IR1     Lvl     AllPub
## 4      4         70 RL           60    9550 Pave   <NA>  IR1     Lvl     AllPub
## 5      5         60 RL           84   14260 Pave   <NA>  IR1     Lvl     AllPub
## 6      6         50 RL           85   14115 Pave   <NA>  IR1     Lvl     AllPub
## 7      7         20 RL           75   10084 Pave   <NA>  Reg     Lvl     AllPub
## 8      8         60 RL           NA   10382 Pave   <NA>  IR1     Lvl     AllPub
## 9      9         50 RM           51    6120 Pave   <NA>  Reg     Lvl     AllPub
## 10    10        190 RL           50    7420 Pave   <NA>  Reg     Lvl     AllPub
## # ... with 1,450 more rows, 71 more variables: LotConfig <chr>,
## #   LandSlope <chr>, Neighborhood <chr>, Condition1 <chr>, Condition2 <chr>,
## #   BldgType <chr>, HouseStyle <chr>, OverallQual <dbl>, OverallCond <dbl>,
## #   YearBuilt <dbl>, YearRemodAdd <dbl>, RoofStyle <chr>, RoofMatl <chr>,
## #   Exterior1st <chr>, Exterior2nd <chr>, MasVnrType <chr>, MasVnrArea <dbl>,
## #   ExterQual <chr>, ExterCond <chr>, Foundation <chr>, BsmtQual <chr>,
## #   BsmtCond <chr>, BsmtExposure <chr>, BsmtFinType1 <chr>, ...
```